UNIVERSITI PUTRA MALAYSIA

RANK-ORDER WEIGHTING OF WEB ATTRIBUTES FOR WEBSITE EVALUATION

MEHRI SAEID

FSKTM 2008 21

RANK-ORDER WEIGHTING OF WEB ATTRIBUTES FOR WEBSITE EVALUATION

By

MEHRI SAEID

Thesis Submitted to the School of Graduate Studies, University Putra Malaysia, in Fulfilment of the Requirement for the Degree of Master of Science

December 2008

To

My Beloved Father and Mother,
My Brothers and Sisters

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfillment of the requirement for the degree of Master of Science

RANK-ORDER WEIGHTING OF WEB ATTRIBUTES FOR WEBSITE EVALUATION

By

MEHRI SAEID

December 2008

Chairman: Associate Professor, Abdul Azim Abd. Ghani, PhD

Faculty: Computer Science and Information Technology

The rapid growth of web applications increases the need to evaluate web applications objectively. In the past few years some works like WebQEM has objectively evaluated the web applications. However, still weighting web attributes which is one step of evaluation of web applications is completely subjective, depending mostly on experts' judgments.

A two-step weighting approach is proposed to solve attribute weighting problem in evaluating web applications in different domains. The approach divides the weighting step into two steps which are ranking and then weighting. Firstly, the web attributes are ranked according to the order of user expectations in web domains, and secondly using rank-order weighting methods (Rank-sum weighting method (RS), Reciprocal of the Ranks weighting method (RR), and Rank-Order Centroid weighting method (ROC)) to elicit weight from the ranked attributes.

A simulation is conducted to compare rank-order weighting methods (RR, RS, and ROC) with the simulated experts. The experts' judgments are simulated in the simulation, assuming that for some particular web attributes, experts weight the attributes completely subjective (randomly without prior ranking). Also for the mentioned attributes, the proposed two-step weighting approach is used.

Two kinds of comparison are done; comparison on weights and comparison on quality scores. Results from simulation are used in comparison to determine which method (RR, RS, and ROC) can be a surrogate for experts' judgments.

From the results of comparison, Rank-sum weighting method (RS) shows 90% of the times completely comply with experts' judgements in terms of rank preservation compared to RR and ROC. This shows that Rank-sum weighting method (RS) is the best method. Rank-sum weighting method (RS) also has very small ValueLoss compared to RR and ROC. From this, it can be said that, using RS weights will give the particular web application a quality score that is not much difference from experts' judgments. Furthermore, 100% of times RS is the best method (compare to RR and ROC) to conform to the experts in terms of choosing the best web application quality. Thus, RS is suggested as a good surrogate for Experts' weights for the attributes when evaluating some web applications.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia
sebagai memenuhi keperluan untuk ijazah Master Sains

PENGUMPUKAN PEMBARAT PENARAFAN-TERTIB ATRIBUT WEB UNTUK PERILAIAN LOKASI WEB

Oleh

MEHRI SAEID

December 2008

Pengerusi: Associate Professor, Abdul Azim Abd. Ghani, PhD

Fakulti: Sains Komputer Dan Teknologi Maklumat

Pertumbuhan pesat aplikasi web menambahkan keperluan untuk menilai aplikasi web secara objektif. Dalam beberapa tahun kebelakangan beberapa usaha seperti WebQEM telah menilai secara objektif aplikasi web. Walau bagaimanapun pengumpukan pemberat kepada atribut web iaitu satu langkah penilaian aplikasi web masih lagi sepenuhnya subjektif bergantung kebanyakannya ke atas pertimbangan pakar.

Satu pendekatan pengumpukan pemberat dua-langkah dicadangkan untuk menyelesaikan masalah pengumpukan pemberat atribut web dalam menilai aplikasi web dalam domain berbeza. Pendekatan ini membahagikan langkah pengumpukan pemberat kepada dua langkah iaitu menyusun kedudukan dan kemudian pengumpukan pemberat. Yang pertama, atribut web disusun kedudukan berdasarkan susunan harapan pengguna dalam domain web, dan yang kedua, menggunakan

kaedah pengumpukan pemberat penarafan-tertib (kaedah pemberatan *Rank-sum* (RS), kaedah pemberatan *Reciprocal of the Ranks* (RR), dan kaedah pemberatan *Rank-Order Centroid* (ROC)) untuk mencungkil pemberat daripada atribut yang tersusun kedudukannya.

Satu simulasi dijalankan untuk membandingkan kaedah pengumpukan pemberat (RR, RS, dan ROC) dengan pakar yang disimulasikan. Pertimbangan pakar disimulasikan dengan anggapan bahawa untuk beberapa atribut web tertentu, pakar mengumpukkan pemberat secara subjektif (secara rawak tanpa penyusunan kedudukan terlebih dahulu).

Dua jenis pembandingan dilakukan; pembandingan ke atas pemberat dan pembandingan ke atas skor kualiti. Keputusan daripada simulasi digunakan dalam perbandingan untuk menentukan kaedah yang dapat menjadi pengganti kepada pertimbangan pakar.

Daripada keputusan perbandingan, kaedah pemberatan *Rank-sum* (RS) menunjukkan 90% mematuhi sepenuhnya pertimbangan pakar dalam mengekalkan susunan kedudukan dibandingkan dengan RR dan ROC. Ini menunjukkan kaedah pemberatan *Rank-sum* (RS) adalah kaedah terbaik. Kaedah pemberatan *Rank-sum* (RS) juga mempunyai *ValueLoss* yang kecil berbanding dengan RR dan ROC. Daripada sini, dapat dinyatakan bahawa menggunakan pemberat RS akan memberi skor kualiti aplikasi web tertentu tidak jauh bezanya dengan pertimbangan pakar. Selain daripada itu, 100% RS adalah kaedah terbaik (dibanding dengan RR dan ROC) mematuhi pakar dalam memilih kualiti terbaik aplikasi web. Sehubungan dengan itu, RS

dicadangkan sebagai pengganti terbaik kepada pemberat pakar untuk atribut apabila menilai beberapa aplikasi web.

ACKNOWLEDGMENTS

First and foremost I would like to express my deep gratefulness to my parent for providing me the opportunity to continue my master's program and financial support. And I'm grateful to my supervisor Associate Professor. Dr. Abdul Azim Abdul Ghani for his kind assistance, critical advice, encouragement, and suggestions during the study and preparation of this thesis. I truly appreciate the time he devoted in advising me and showing me the proper directions to continue this research.

I would also like to express my gratitude to my co-supervisor Associate Professor Hassan Selamat, to whom I'm grateful for his practical experience and knowledge that made an invaluable contribution to this thesis.

I also owe thanks to the people who provided assistance and gave advice. Last but not the least the deepest appreciation goes to my fiancé Hamid for his contentious support and encouragement. Another thank you also goes to Mrs. Hazura for her help and support.

I certify that an Examination Committee has met on **date of viva** to conduct the final examination of **Mehri Saeid** on her **Master of Science** thesis entitled " **RANK-ORDER WIEGHTING OF WEB ATTRIBUTES FOR WEBSITE EVALUATION**" in accordance with Universiti Pertanian Malaysia (Higher Degree) Act 1980 and Universiti Pertanian Malaysia (Higher Degree) Regulations 1981. The Committee recommends that the candidate be awarded the relevant degree. Members of the Examination Committee are as follows:

Ali Mamat, PhD
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Chairman)

Rusli Abdullah, PhD
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Internal Examiner)

Hamidah Ibrahim, PhD
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Internal Examiner)

Safaai Deris, PhD
Faculty of Computer Science and Information System
Universiti Technology Malaysia
(External Examiner)

BUJANG KIM HUAT, PhD
Professor and Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia
Date: 19 February 2009

This thesis submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfillment of the requirements for the degree of Master of Science. Members of the Supervisory Committee are as follows:

Abdul Azim Abdul. Ghani, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Chairman)

Mohd Hassan b Selamat, PhD
Lecturer
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

HASANAH MOHD GHAZALI,PHD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 9 April 2009

DECLARATION

I hereby declare that the thesis is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at UPM or other institutions.

Mehri Saeid
Date:

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATION

AHP	Analytic Hierarchy Process
AvgRSQ	Average quality scores using RS weights
AvgROCQ	Average quality scores using ROC weights
AvgRRQ	Average quality scores using RR weights
AvgExprt	Average quality scores using TRUE weights
Decision Makers	DM
DMW	Decision Makers' Weights Matrix
ISO	International Organization of Standardization
IS	Information System
HCI	Human Computer Interaction
MDMW	Mean of Decision Makers' Weights Matrix
RS	Rank-sum weighting method
RR	Reciprocal of the Ranks weighting method
ROC	Rank-Order Centroid weighting method
ROCW	ROC Weight Matrix
RSW	RS Weight Matrix
RRW	RR Weight Matrix
SVM	Simulated Value Matrix
VL	ValueLoss
Web-App	Web Application
WE	Web Engineering
TRUE Weights	Weights assigned by Decision Makers (Here Web application Evaluator or Experts)

CHAPTER 1

INTRODUCTION

1.1 Background

Some researchers consider that quality of product or service is what the end-user or customer receives from it, not what the provider or seller put into it. Hence, a website should satisfy its customers' needs to ensure repeat their visits, and achieve their loyalty.

In order to evaluate the quality of a website, a number of attempts at evaluation of consumer-oriented websites has been developed and published in the last few years. Some of them were in a purely subjective form of individual preferences of the assessor, and some were in the objective form of statistical measurement, such as monitoring the download time of the site and site traffics (Hung & McQueen, 2004).

Those who assess websites are called evaluator here. They assess the "work" of web designers and developers, investigating whether they have created a website that serves its purpose. More specifically they examine whether the website provides the customers with means to effectively interact with the company, motivated them to purchase the product and service they were looking for and make their visit so easy and enjoyable that they would like to return and visit it again (Zhang & Dran, 2001).

Actually what happened in the evaluation of a website is that the web evaluator estimates the quality of a web application upon some specific features. These features are called attributes in this research. It is clear that any type of web

application in terms of some particular attributes can be evaluated, as Olsina mentioned, many attributes can be reused among different web application domains (Olsina *et al.*, 2000). However one attribute has different importance in different web application domains. This means for example security attributes in e-commerce domain are the most important one but not in an entertainment web application domain. So what the evaluator should do is to consider these differences of importance some how in the evaluation process. Usually evaluators weigh the attributes in terms of the attributes' importance in order to bring this importance to evaluation of web application. However understanding the importance and consequently the weights is not easy. The evaluators should have good experiences to know which attribute is more important than the other attribute in the particular domain, and after understanding that, s/he may be able to weigh the attribute.

None of the above tasks; understanding the attribute's importance and weighting it, is easy. In previous works on evaluation of website, noting is mentioned explicitly about weighting web attributes. There are quality evaluation models that group and classify web attributes to ease the web quality evaluation in a particular domain. However nothing has been said about how the attributes have been weighted. Most of them rely just on experts' experiences and judgements. The weights of the attributes are fixed in the quality model proposed by Olsina *et al.* (2002) which are from their experiences from previous projects. There is no process about how they weighted the attributes. This can bring difficulties to an evaluator in evaluating web applications in domains other than those domains of the predefined quality models. These difficulties are because that the importance of attributes may be different

among various domains. Consequently weighting of the attributes could not be done precisely.

On the other hand, Zhang *et al.* (2001) have done a valuable research on features of users' satisfaction in some web application domains. In their research, they compared websites in six web application domains upon fourteen clusters or families of features. The six application domains they used are finance, education, entertainment, e-commerce, government and medicine. They have ranked the clusters of attributes in different domains, based on the user satisfaction and expectations.

It seems looking at web attribute weighting as a weighting decision problem and consequently solving this problem by providing objective ways for it, directs us to replace current subjective web attribute weighting by objective ranking methods such as rank-order-weighting, as well as using Zhang's results in ranking attributes in the web domain of web application in order to use them in web quality evaluation.

We believe that from researches like Zhang *et al.* (2001), an evaluator can understand the importance' of attributes, so by having the importance, the ranks of the attributes in a domain exists, which can be used to weight the attributes.

1.2 Problem Statement

Recently some researches have been conducted to make web quality evaluation quantitative and less subjective (Olsina & Rassi, 2002; Olsina et. al. 2007). Their work is a great step in quantitative web evaluation. One task in quantitative web

evaluation is attribute weighting. Still it is completely subjective, depending on domain experts' experiences to directly weight the attributes. Thus, it is not easy for a web evaluator to elicit weights for attribute according to different web application domains and also in one domain but among all attributes in the domain. On the other hand, from literature, it can be concluded that usually ranking is easier than weighting for non expert or even experts (Moshkovich *et al.*, 2001). In general, in any weighting decision problem, using experts judgements to directly assign weights is a problem (Barron & Barrett, 1996; Ahn & Park 2006). These problems also can be recognized in weighting web attributes in web quality evaluation process. The necessity of this research can be mentioned as:

1. The web quality evaluator may be unavailable, unable, or unwilling to specify sufficiently precise weights; or

2. There may be no single domain expert, and the evaluator group may not even be able to agree on a ranking of attributes in one web application domain, and also not to directly assign weights.

3. Subjectivity doesn't have repeated measurement, that means if repeating a subjective measurement its not guarantied that the same result will be achieved.

1.3 Research Objective

The main objective of this research is to propose a weighting approach, in order to use it as the surrogate for subjective expert weighting of web applications' attributes in web quality evaluation process.

1.4 Research Scope

This research is scoped according to the following delimitations:

- Each attribute must be classified just in one cluster or category; this is because ranking of the attributes depends on what cluster the attribute is assigned to, so the attribute should belong to one cluster.

- Website should be correctly assigned to a web application domain; this is also because the rank of attributes is different in each web domain, so in order to obtain the right rank for the attribute in the domain, the web application should be assigned to the correct web domain.

- Number of attributes is better to be less than 7 in one group, this is because, having more than 7 attributes in a group for weighting, and using rank-order weighting formula, cause some of them to get very small weights; less than 0.01. This low weight reduces the effect of the attribute very much which is not intended here.

- Attributes belonging to the same cluster have the same importance, and consequently same weights, this is because attributes that are classified in the same cluster, would have the same rank, so it is supposed to have the same weights too.

- Attributes used in simulation have no dependency on each other, they are independent; this means that it is assumed that if one attribute has high quality score it does not cause another attribute having less influence or vice versa.

- The sum of weights of the attributes must be 1: $\sum_{i=1}^{n} w_i = 1$. The value of each attribute is between 0-100. Also the final quality score for the obligatory websites is expressed in percent (scale of 100). So having the total weights of attributes as unit (=1), and multiplying each attribute's measured value (between 0-100) by the attribute's weight (less than 1), will result the final quality score of the website in a number between 0-100 (it is expressed in percent).

1.5 Thesis Organization

This thesis is outlined in six chapters. This chapter provides background information about web evaluation and web attribute weighting, and explains the problem statement. The objective of this research is also included in this chapter. Chapter 2 consists of the reviewed literature of the related works. Chapter 3 contains a general description of research methodology, and also the criteria which upon them we will evaluate our work. Chapter 4 consists of detailed steps of simulation study. In chapter 5, we bring the results of simulation. The performance of proposed method is evaluated also. Chapter 6 shows the conclusion that summarizes the most important aspects of research, and ends with contribution and suggested future works.

CHAPTER 2

LITERATURE REVIEW

2.1. Introduction

This chapter is divided into ten main sections. Section 2.2 provides the discussion about web application and the concept of quality and quality of web application. Section 2.3 discusses about web quality model. Section 2.4 provides a brief discussion about evaluation process. Section 2.5 discusses about attribute weighting in web application. Section 2.6 is about different ranking methods. Section 2.7 provides the weighting methods categories. Section 2.8 discusses the evaluation of the weighting methods. A brief drawback of pervious works is described in Section 2.9, and finally Section 2.10 summarizes the literature review.

2.2. Web Application and its Quality

Within a short period, the internet and World Wide Web have become ubiquitous, surpassing all other technological developments in our history. They have also grown rapidly in their scope and extent of use, significantly affecting all aspects of our lives. Industries such as manufacturing, travel and hospitality, banking, education, and government are web-enabled to improve and enhance their operations.

E-Commerce has expanded quickly, cutting across national boundaries. Even traditional legacy information and database systems have migrated to the web. As a result, we increasingly depend on a range of web applications (Ginige & Murugesan, 2001), but unfortunately as Nielsen (2001) stated in his website; most websites are guilty of poor quality and low robustness, furthermore he stated that, the main goal